



SEQUENCE LISTING

GENERAL INFORMATION:

- (i) APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi
Chen, Wei Ning
 - (ii) TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF
 - (iii) NUMBER OF SEQUENCES: 11
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Ladas & Parry
 - (B) STREET: 26 West 61 Street
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 10023
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/719,528
 - (B) FILING DATE: 19-DEC-2000
 - (C) CLASSIFICATION: 435
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/SG98/00046
 - (B) FILING DATE: 19-JAN-1998
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Mass, Clifford J.
 - (B) REGISTRATION NUMBER: 30,086
 - (C) REFERENCE/DOCKET NUMBER: U-014987-0
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 708-1800
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3215 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACAACA TTCCACCAAG CTCTGCTAGA TCCCAGGGTG AGGGGCCTAT ATTTTCCTGC	60
TGGTGGCTCC AGTTCCGGAA CAGTAAACCC TGTTCCGACT ACTGCCTCTC CCATATCGTC	120
AATCTTCTCG AGGACTGGGG ACCCTGCACC GAACATGGAG AACACAACAT CAGGATTCCCT	180
AGGACCCCTG CTCGTGTTAC AGGCGGGGTT TTTCTCGTTG ACAAGAATCC TCACAATACC	240
GCAGAGTCTA GACTCTGGTG GACTTCTCTC AATTTTCTAG GGGGAGCACC CACGTGTTCC	300
TGGCCAAAT TCGCAGTCCC CAACCTCCAA TCACTACCA ACCTCTTGTC CTCCAATTTG	360
TCCTGGCTAT CGCTGGATGT GTCTGCGGCG TTTTATCATA TTCCTCTTCA TCCTGCTGCT	420
ATGCCTCATC TTCTTGTTGG TTCTTCTGGA CTACCAAGGT ATGTTGCCCG TTTGTCTCT	480
ACTTCCAGGA ACATCAACCA CCAGCACGGG GCCATGCAAG ACCTGCACGA CTCCTGCTCA	540
AGGAACTCT ACGTTTCCCT CTTGTTGCTG TACAAAACCT TCGGACGGAA ACTGCACTTG	600
TATTTCCATC CCATCATCCT GGGCTTTCGC AAGATTCCCTA TGGGAGTGGG CCTCAGTCCG	660
TTTCTCTGG CTCAGTTTAC TAGTGCCATT TGTTCACTGG TTCGTAGGGC TTTCCCCAC	720
TGTTTGGCTT TCAGTTATAT GGATGATGTG GTATTGGGGG CGAAGTCTGT ACAACATCTT	780
GAGTCCCTTT TTACCTCTAT TACCAATTTT CTTTTGTCTT TGGGTATACA TTTAAACCT	840
AATAAAACCA AACGTTGGGG CTACTCCCTT AACTTCATGG GATATGTAAT TGGAAGTTGG	900
GGTACTTTAC CGCAGGAACA TATTGTACTA AAACCTCAAGC AATGTTTTCG AAAACTGCCT	960
GTAAATAGAC CTATTGATTG GAAAGTATGT CAAAGAATTG TGGGTCTTTT GGGCTTTGCT	1020
GCCCCTTTTA CACAATGTGG CTATCCTGCC TTGATGCCTT TATATGCATG TATACAATCT	1080
AAGCAGGCTT TCACTTTCTC GCCAACTTAC AAGGCCTTTC TGTGTAAACA ATATCTGAAC	1140
CTTTACCCCG TTGCCCGGCA ACGGTCCGGT CTCTGCCAAG TGTTTGCTGA CGCAACCCCC	1200
ACTGGATGGG GCTTGGCCAT AGGCCATCAG CGCATGGCTG GAACCTTTCT GGCTCCTCTG	1260
CCGATCCATA CTGCGGAACCT CCTAGCAGCT TGTTTTGCTC GCAGCCGGTC TGGAGCAAAA	1320
CTTATCGGAA CCGACAACCTC TGTTGTCCTC TCTCGGAAAT ACACCTCCTT TCCATGGCTG	1380
CTAGGGTGTG CTGCCAACTG GATCCTGCGC GGGACGTCCT TTGTCTACGT CCCGTCGGCG	1440
CTGAATCCCG CGGACGACCC GTCTCGGGGC CGTTTGGGGC TCTACCGTCC CCTTCTTCAT	1500
CTGCCGTTCC GGCCGACCAC GGGGCGCACC TCTCTTTACG CGGTCTCCCC GTATGTGCCT	1560
TCTCATCTGC CGGACCGTGT GCACTTCGCT TCACCTCTGC ACGTCGCATG GAGACCACCG	1620
TGAACGCACG CCAGGTCTTG CCCAAGGTCT TATATAAGAG GACTCTTGGA CTCTCAGCAA	1680
TGTCAACGAC CGACCTTGAG GCATACTTCA AAGACTGTGT GTTTAAAGAC TGGGAGGAGT	1740

TGGGGGAGGA	GATTAGGTTA	AAGATTTATG	TACTAGGAGG	CTGTAGGCAT	AAATTGGTCT	1800
GTTCACCAGC	ACCATGCAAC	TTTTTCTCCT	CTGCCTAATC	ATCTCATGTT	CATGTCCTAC	1860
TGTTCAAGCC	TCCAAGCTGT	GCCTTGGGTG	GCTTTGGGAC	ATGGACATTG	ACCCGTATAA	1920
AGAATTTGGA	GCATCTGCTG	AGTTACTCTC	TTTTTTGCCT	TCTGACTTCT	TTCCGTCTAT	1980
TCGAGATCTC	CTCGACACCG	CCTCTGCTCT	GTATCGGGAG	GCCTTAGAGT	CTCCGGAACA	2040
TTGTTGCGCT	CACCATACAG	CACTCAGGCA	AGCTATTTTG	TGTTGGGGTG	AGTTGATGAA	2100
TCTGGCCACC	TGGGTGGGAA	GTAATTTGGA	AGATCCAGCA	TCCAGGGAAT	TAGTAGTCAG	2160
CTATGTCAAC	GTTAATATGG	GCCTAAAACT	CAGACAAATA	TTGTGGTTTC	ACATTTCTTG	2220
TCTTACTTTT	GGAAGAGAAA	CTGTTCTTGA	GTACTTGGTA	TCTTTTGGAG	TGTGGATTCTG	2280
CACTCCTACC	GCTTACAGAC	CACCAAATGC	CCCTATCTTA	TCAACACTTC	CGGAAACTAC	2340
TGTTGTTAGA	CGACGAGGCA	GGTCCCCTAG	AAGAAGAACT	CCCTCGCCTC	GCAGACGAAG	2400
GTCTCAATCG	CCGCGTCGCA	GAAGATCTCA	ATCTCGGGAA	TCTCAACGTT	AGTATTCCTT	2460
GGACTCATAA	GGTGGGAAAC	TTTACTGGGC	TTTATTCTTC	TACTGTACCT	GTCTTTAATC	2520
CCGAGTGGCA	AATTCCTTCC	TTTCCTCACA	TTCATTTACA	AGAGGACATT	ATTAATAGAT	2580
GTCAACAATA	TGTGGGCCCT	CTTACAGTTA	ATGAAAAAAG	AAGATTAAAA	TTAATTATGC	2640
CTGCTAGGTT	TTATCCTAAC	CTTACTAAAT	ATTTGCCCTT	AGACAAAGGC	ATTAAACCGT	2700
ATTATCCTGA	ACATGCAGTT	AATCATTACT	TCAAAACTAG	GCATTATTTA	CATACTCTGT	2760
GGAAGGCTGG	CATTCTATAT	AAGAGAGAAA	CTACACGCAG	CGCCTCATTT	TGTGGGTCAC	2820
CATATTCTTG	GGAACAAGAG	CTACAGCATG	GGAGGTTGGT	CTTCCAAACC	TCGACAAGGC	2880
ATGGGGAGCA	ATCTTGCTGT	TCCCAATCCT	CTGGGATTCT	TTCCCGATCA	CCAGTTGGAC	2940
CCTGCGTTTCG	GAGCCAACTC	AAACAATCCA	GATTGGGACT	TCAACCCCAA	CAAGGATCAC	3000
TGGCCAGAGG	CAAATCAGGT	AGGAGTGGGA	GCATTTCGGGC	CAGGGTTCAC	CCCACCACAC	3060
GGCGGTCTTT	TGGGGGGGAG	CCCTCAGGCT	CAGGGCATAT	TGACAACAGT	GCCAGCAGCA	3120
CCTCCTCCTG	CCTCCACCAA	TCGGCAGTCA	GGAAGACAGC	CTACTCCCAT	CTCTCCACCT	3180
CTAAGAGACA	GTCATCCTCA	GGCCACGCAG	TGGAA			3215

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 843 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Leu	Ser	Tyr	Gln	His	Phe	Arg	Lys	Leu	Leu	Leu	Leu	Asp	Asp	
1			5						10					15		
Glu	Ala	Gly	Pro	Leu	Glu	Glu	Glu	Leu	Pro	Arg	Leu	Ala	Asp	Glu	Gly	
			20					25					30			
Leu	Asn	Arg	Arg	Val	Ala	Glu	Asp	Leu	Asn	Leu	Gly	Asn	Leu	Asn	Val	
			35				40					45				
Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	Phe	Thr	Gly	Leu	Tyr	Ser	
	50					55					60					
Ser	Thr	Val	Pro	Val	Phe	Asn	Pro	Glu	Trp	Gln	Ile	Pro	Ser	Phe	Pro	
65					70					75					80	
His	Ile	His	Leu	Gln	Glu	Asp	Ile	Ile	Asn	Arg	Cys	Gln	Gln	Tyr	Val	
			85						90					95		
Gly	Pro	Leu	Thr	Val	Asn	Glu	Lys	Arg	Arg	Leu	Lys	Leu	Ile	Met	Pro	
			100					105						110		
Ala	Arg	Phe	Tyr	Pro	Asn	Leu	Thr	Lys	Tyr	Leu	Pro	Leu	Asp	Lys	Gly	
		115					120						125			
Ile	Lys	Pro	Tyr	Tyr	Pro	Glu	His	Ala	Val	Asn	His	Tyr	Phe	Lys	Thr	
	130					135					140					
Arg	His	Tyr	Leu	His	Thr	Leu	Trp	Lys	Ala	Gly	Ile	Leu	Tyr	Lys	Arg	
145					150					155					160	
Glu	Thr	Thr	Arg	Ser	Ala	Ser	Phe	Cys	Gly	Ser	Pro	Tyr	Ser	Trp	Glu	
				165					170					175		
Gln	Glu	Leu	Gln	His	Gly	Arg	Leu	Val	Phe	Gln	Thr	Ser	Thr	Arg	His	
			180					185					190			
Gly	Asp	Glu	Ser	Cys	Cys	Ser	Gln	Ser	Ser	Gly	Ile	Leu	Ser	Arg	Ser	
		195					200					205				
Pro	Val	Gly	Pro	Cys	Val	Arg	Ser	Gln	Leu	Lys	Gln	Ser	Arg	Leu	Gly	
	210					215					220					
Leu	Gln	Pro	Gln	Gln	Gly	Ser	Leu	Ala	Arg	Gly	Lys	Ser	Gly	Arg	Ser	
225					230					235					240	
Gly	Ser	Ile	Arg	Ala	Arg	Val	His	Pro	Thr	Thr	Arg	Arg	Ser	Phe	Gly	
				245					250					255		
Gly	Glu	Pro	Ser	Gly	Ser	Gly	His	Ile	Asp	Asn	Ser	Ala	Ser	Ser	Thr	
			260					265					270			

Ser Ser Cys Leu His Gln Ser Ala Val Arg Lys Thr Ala Tyr Ser His
 275 280 285
 Leu Ser Thr Ser Lys Arg Gln Ser Ser Ser Gly His Ala Val Glu Leu
 290 300
 His Asn Ile Pro Pro Ser Ser Ala Arg Ser Gln Gly Glu Gly Pro Ile
 305 310 315 320
 Phe Ser Cys Trp Trp Leu Gln Phe Arg Asn Ser Lys Pro Cys Ser Asp
 325 330 335
 Tyr Cys Leu Ser His Ile Val Asn Leu Leu Glu Asp Trp Gly Pro Cys
 340 345 350
 Thr Glu His Gly Glu His Asn Ile Arg Ile Pro Arg Thr Pro Ala Arg
 355 360 365
 Val Thr Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala
 370 375 380
 Glu Ser Arg Leu Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ala Pro
 385 390 395 400
 Thr Cys Ser Trp Pro Lys Phe Ala Val Pro Asn Leu Gln Ser Leu Thr
 405 410 415
 Asn Leu Leu Ser Ser Asn Leu Ser Trp Leu Ser Leu Asp Val Ser Ala
 420 425 430
 Ala Phe Tyr His Ile Pro Leu His Pro Ala Ala Met Pro His Leu Leu
 435 440 445
 Val Gly Ser Ser Gly Leu Pro Arg Tyr Val Ala Arg Leu Ser Ser Thr
 450 455 460
 Ser Arg Asn Ile Asn His Gln His Gly Ala Met Gln Asp Leu His Asp
 465 470 475 480
 Ser Cys Ser Arg Lys Leu Tyr Val Ser Leu Leu Leu Leu Tyr Lys Thr
 485 490 495
 Phe Gly Arg Lys Leu His Leu Tyr Ser His Pro Ile Ile Leu Gly Phe
 500 505 510
 Arg Lys Ile Pro Met Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln
 515 520 525
 Phe Thr Ser Ala Ile Cys Ser Val Val Arg Arg Ala Phe Pro His Cys
 530 535 540
 Leu Ala Phe Ser Tyr Met Asp Asp Val Val Leu Gly Ala Lys Ser Val
 545 550 555 560
 Gln His Leu Glu Ser Leu Phe Thr Ser Ile Thr Asn Phe Leu Leu Ser
 565 570 575

Leu	Gly	Ile	His	Leu	Asn	Pro	Asn	Lys	Thr	Lys	Arg	Trp	Gly	Tyr	Ser		
			580					585					590				
Leu	Asn	Phe	Met	Gly	Tyr	Val	Ile	Gly	Ser	Trp	Gly	Thr	Leu	Pro	Gln		
		595					600					605					
Glu	His	Ile	Val	Leu	Lys	Leu	Lys	Gln	Cys	Phe	Arg	Lys	Leu	Pro	Val		
	610					615					620						
Asn	Arg	Pro	Ile	Asp	Trp	Lys	Val	Cys	Gln	Arg	Ile	Val	Gly	Leu	Leu		
625					630					635					640		
Gly	Phe	Ala	Ala	Pro	Phe	Thr	Gln	Cys	Gly	Tyr	Pro	Ala	Leu	Met	Pro		
				645					650					655			
Leu	Tyr	Ala	Cys	Ile	Gln	Ser	Lys	Gln	Ala	Phe	Thr	Phe	Ser	Pro	Thr		
			660					665					670				
Tyr	Lys	Ala	Phe	Leu	Cys	Lys	Gln	Tyr	Leu	Asn	Leu	Tyr	Pro	Val	Ala		
	675						680					685					
Arg	Gln	Arg	Ser	Gly	Leu	Cys	Gln	Val	Phe	Ala	Asp	Ala	Thr	Pro	Thr		
	690					695					700						
Gly	Trp	Gly	Leu	Ala	Ile	Gly	His	Gln	Arg	Met	Ala	Gly	Thr	Phe	Leu		
705					710					715					720		
Ala	Pro	Leu	Pro	Ile	His	Thr	Ala	Glu	Leu	Leu	Ala	Ala	Cys	Phe	Ala		
				725				730						735			
Arg	Ser	Arg	Ser	Gly	Ala	Lys	Leu	Ile	Gly	Thr	Asp	Asn	Ser	Val	Val		
			740					745					750				
Leu	Ser	Arg	Lys	Tyr	Thr	Ser	Phe	Pro	Trp	Leu	Leu	Gly	Cys	Ala	Ala		
	755						760					765					
Asn	Trp	Ile	Leu	Arg	Gly	Thr	Ser	Phe	Val	Tyr	Val	Pro	Ser	Ala	Leu		
	770					775					780						
Asn	Pro	Ala	Asp	Asp	Pro	Ser	Arg	Gly	Arg	Leu	Gly	Leu	Tyr	Arg	Pro		
785					790					795					800		
Leu	Leu	His	Leu	Pro	Phe	Arg	Pro	Thr	Thr	Gly	Arg	Thr	Ser	Leu	Tyr		
				805					810					815			
Ala	Val	Ser	Pro	Tyr	Val	Pro	Ser	His	Leu	Pro	Asp	Arg	Val	His	Phe		
			820					825					830				
Ala	Ser	Pro	Leu	His	Val	Ala	Trp	Arg	Pro	Pro							
	835						840										

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Gly	Gly	Trp	Ser	Ser	Lys	Pro	Arg	Gln	Gly	Met	Gly	Thr	Asn	Leu	1	5	10	15
Ala	Val	Pro	Asn	Pro	Leu	Gly	Phe	Phe	Pro	Asp	His	Gln	Leu	Asp	Pro	20	25	30	
Ala	Phe	Gly	Ala	Asn	Ser	Asn	Asn	Pro	Asp	Trp	Asp	Phe	Asn	Pro	Asn	35	40	45	
Lys	Asp	His	Trp	Pro	Glu	Ala	Asn	Gln	Val	Gly	Val	Gly	Ala	Phe	Gly	50	55	60	
Pro	Gly	Phe	Thr	Pro	Pro	His	Gly	Gly	Leu	Leu	Gly	Gly	Ser	Pro	Gln	65	70	75	80
Ala	Gln	Gly	Ile	Leu	Thr	Thr	Val	Pro	Ala	Ala	Pro	Pro	Pro	Ala	Ser	85	90	95	
Thr	Asn	Arg	Gln	Ser	Gly	Arg	Gln	Pro	Thr	Pro	Ile	Ser	Pro	Pro	Leu	100	105	110	
Arg	Asp	Ser	His	Pro	Gln	Ala	Thr	Gln	Trp	Asn	Ser	Thr	Thr	Phe	His	115	120	125	
Gln	Ala	Leu	Leu	Asp	Pro	Arg	Val	Arg	Gly	Leu	Tyr	Phe	Pro	Ala	Gly	130	135	140	
Gly	Ser	Ser	Ser	Gly	Thr	Val	Asn	Pro	Val	Pro	Thr	Thr	Ala	Ser	Pro	145	150	155	160
Ile	Ser	Ser	Ile	Phe	Ser	Arg	Thr	Gly	Asp	Pro	Ala	Pro	Asn	Met	Glu	165	170	175	
Asn	Thr	Thr	Ser	Gly	Phe	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	Ala	Gly	180	185	190	
Phe	Phe	Ser	Leu	Thr	Arg	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	Ser	195	200	205	
Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Ala	Pro	Thr	Cys	Pro	Gly	210	215	220	
Gln	Asn	Ser	Gln	Ser	Pro	Thr	Ser	Asn	His	Ser	Pro	Thr	Ser	Cys	Pro	225	230	235	240
Pro	Ile	Cys	Pro	Gly	Tyr	Arg	Trp	Asn	Cys	Leu	Arg	Arg	Phe	Ile	Ile	245	250	255	
Phe	Leu	Phe	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu	260	265	270	

Asp	Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Leu	Pro	Gly	Thr	Ser	275	280	285	
Thr	Thr	Ser	Thr	Gly	Pro	Cys	Lys	Thr	Cys	Thr	Thr	Pro	Ala	Gln	Gly	290	295	300	
Asn	Ser	Thr	Phe	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Ser	Asp	Gly	Asn	305	310	315	320
Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Ala	Arg	Phe	Leu	325	330	335	
Trp	Glu	Trp	Ala	Ser	Val	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro	340	345	350	
Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val	355	360	365	
Ile	Trp	Met	Met	Trp	Tyr	Trp	Gly	Arg	Ser	Leu	Tyr	Asn	Ile	Leu	Ser	370	375	380	
Pro	Phe	Leu	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	Ile	385	390	395	400

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 212 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Gln	Leu	Phe	Leu	Leu	Cys	Leu	Ile	Ile	Ser	Cys	Ser	Cys	Pro	Thr	1	5	10	15
Val	Gln	Ala	Ser	Lys	Leu	Cys	Leu	Gly	Trp	Leu	Trp	Asp	Met	Asp	Ile	20	25	30	
Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Ser	Ala	Glu	Leu	Leu	Ser	Phe	Leu	35	40	45	
Pro	Ser	Asp	Phe	Phe	Pro	Ser	Ile	Arg	Asp	Leu	Leu	Asp	Thr	Ala	Ser	50	55	60	
Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	Ser	Pro	His	65	70	75	80
His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Cys	Trp	Gly	Glu	Leu	Met	Asn	85	90	95	
Leu	Ala	Thr	Trp	Val	Gly	Ser	Asn	Leu	Glu	Asp	Pro	Ala	Ser	Arg	Glu	100	105	110	

Leu Val Val Ser Tyr Val Asn Val Asn Met Gly Leu Lys Leu Arg Gln
 115 120 125
 Ile Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val
 130 135 140
 Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Thr Ala
 145 150 155 160
 Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr
 165 170 175
 Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser Pro
 180 185 190
 Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser Gln Ser Arg
 195 200 205
 Glu Ser Gln Arg
 210

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala Ala Arg Val Cys Cys Gln Leu Asp Pro Ala Arg Asp Val Leu
 1 5 10 15
 Cys Leu Arg Pro Val Gly Ala Glu Ser Arg Gly Arg Pro Val Ser Gly
 20 25 30
 Pro Phe Gly Ala Leu Pro Ser Pro Ser Ser Ser Ala Val Pro Ala Asp
 35 40 45
 His Gly Ala His Leu Ser Leu Arg Gly Leu Pro Val Cys Ala Phe Ser
 50 55 60
 Ser Ala Gly Pro Cys Ala Leu Arg Phe Thr Ser Ala Arg Arg Met Glu
 65 70 75 80
 Thr Thr Val Asn Ala Arg Gln Val Leu Pro Lys Val Leu Tyr Lys Arg
 85 90 95
 Thr Leu Gly Leu Ser Ala Met Ser Thr Thr Asp Leu Glu Ala Tyr Phe
 100 105 110
 Lys Asp Cys Val Phe Lys Asp Trp Glu Glu Leu Gly Glu Glu Ile Arg
 115 120 125

Leu Lys Ile Tyr Val Leu Gly Gly Cys Arg His Lys Leu Val Cys Ser
 130 135 140
 Pro Ala Pro Cys Asn Phe Phe Ser Ser Ala
 145 150

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAAGCTTAT GCCCCTATCT TATCAACACT TCCGGA

36

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGTCTAGAC TCTGCGGTAT TGTGA

25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGTCTAGAC TCGTGGTGGA CTTCT

25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAGAATTCT CACGGTGGTC TCCATGCGAC GT

32

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTGTTTACG TCCCGT

16

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATAAGCTTAT GCCCCTATCT TATCAACACT TCCGGA

36